This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2707103, IEEE Transactions on Smart Grid

Distributed Coordination of EV Charging with Renewable Energy in a Microgrid of Buildings

Yu Yang, *Student Member, IEEE,* Qing-Shan Jia, *Senior Member, IEEE,* Geert Deconinck, *Senior Member, IEEE,* Xiaohong Guan, *Fellow, IEEE,* Zhifeng Qiu, *Member, IEEE,* and Zechun Hu, *Member, IEEE*

***Abstract*—With the rapid development of electric vehicles (EVs), the consequent charging demand represents a significant new load on the power grids. The huge number of high-rise buildings in big cities and modern technological advances have created conditions to mount on-site wind power generators on the buildings. Since modern buildings are usually equipped with large parking lots for EVs, it shows vital practical significance to utilize the on-site wind power generation to charge EVs parked in the buildings. In this paper, we first use a case study in Beijing to show that the on-site wind power generation of high-rise buildings can potentially support all the EVs in the city. Considering that the charging demand of EVs usually doesn't align with the uncertain wind power, the coordination of EV charging with the locally generated wind power in a microgrid of buildings is investigated and three main contributions are made. *First*, we investigate the problem and formulate it as a Markov decision process (MDP), which incorporates the random driving requirements of EVs among the buildings. *Second*, we develop a distributed simulation-based policy improvement (DS-BPI) method, which can improve from heuristic and experience-based policies. *Third*, the performance of the distributed policy improvement method is proved. We compare DSBPI with a central version method on two case studies. The DSBPI method demonstrates good performance and scalability.**

***Index Terms*—Electric vehicles (EVs), building mounted wind power, Markov decision process, distributed optimization**

I. Introduction

In recent years, electric vehicles (EVs) have attracted extensive concerns with the incentives to decrease carbon emissions and alleviate fuel depletion. However, with the popularity of EVs, the consequent charging demand represents a significant new load, producing non-negligible negative impacts on the power grids [1]. This has been recognized as one of the most challenging issues that hinder the wide-adoption of EVs around the world [2].

This work is supported in part by the National Key Research and Development Program of China (2016YFB0901905), the National Natural Science Foundation of China under grants (No. 61673229, 61174072, 61222302, 91224008, 61221063, 61403429, 61621062, and U1301254), TNList Founding for Excellent Young Scholar, 111 International Collaboration Project of China (No. B06002), the Program for New Star of Science and Technology in Beijing (No. xx2014B056), Belgium Science Policy Office (F-PM2015/ZKD0579), and the ISP/13/05TS from University of Leuven.

Y. Yang, Q.-S. Jia, and X. Guan are with the Center for Intelligent and Networked Systems (CFINS), Department of Automation, Tsinghua University, Beijing 100084, China (E-mail:{ yangyu13@mails., jiaqs@, xhguan@ } tsinghua.edu.cn). X. Guan is also with the Ministry of Education Key Lab for Intelligent Networks and Network Security Laboratory, Xi'an Jiaotong University, Xi'an, Shaanxi Province, 710049, China.

G. Deconinck, and Z. Qiu are with the Department of Electrical Engineering, Electrical Energy, EnergyVille, University of Leuven, Leuven 3001, Belgium. Z. Qiu is also with the Department of Electrical Engineering, Central South University, Changsha, Hunan Province, 410000, China.

Z. Hu is with the Department of Electrical Engineering, Tsinghua University, Beijing, 100084, China.

Q.-S. Jia is the corresponding author.

Currently, the huge number of high-rise buildings and modern technological advances have created conditions to mount on-site wind power generators on buildings. The reported total number of buildings higher than 200 meters has exceeded $1,000$ in 2015 around the world, while China accounts for about 40% [3]. The rapid increase of high-rise buildings, especially in cities, shows great potential to fully explore wind power of buildings due to less turbulence and abundant wind resources [4]. Meanwhile, much effort has been devoted to the development of wind turbines for buildings over the past decades [5], [6], [7]. For instance, a newly designed wind turbine for buildings, which is both silent and efficient at converting wind into energy, shows great promise to be widely adopted [8]. Besides, the axial load of wind turbines on buildings has been halved in the last few years [9]. Since modern buildings, including residential, office and commercial buildings, are usually equipped with large parking lots for EVs, the issue to utilize on-site wind power of buildings to charge the EVs parked in the buildings shows vital practical significance in various aspects. *First*, the huge number of high-rise buildings in cities provides enormous wind energy to be fully explored to support EV charging, thus decreasing their impacts on the power grids. *Second*, unlike remote wind farms, the on-site wind power of buildings can be locally utilized to charge EVs parked there without a need to build a new high-voltage transmission system. *Third*, the variability of wind power can be regulated by properly scheduling EV charging due to their charging flexibility.

In the literature, much effort has been devoted to the management of EV charging with various attempts. For example, Sundström *et al.* [10] established an individual charging plan for each vehicle of an aggregator to avoid grid congestion. Vandael *et al.* [11] proposed a three-step approach to manage the charging load of massive plug-in electric vehicles (PHEVs) to minimize the cost for energy supplier. Besides, with the popularity of renewable energy, the coordination of EV charging with renewable energy has attracted extensive attention. For instance, Wu *et at.* [12] developed three heuristic dispatching approaches for EVs to improve the matching of energy consumption and wind supply over the night. Kou *et al.* [13] proposed a two-layer coordination of PEV charging with wind power based on the prediction of wind power and EV charging demand in each predefined planning horizon over the off-peak time. Similarly, Guo *et al.* [15] addressed a two-stage framework for the economic operation of an EV parking deck with on-site renewable generation using the same idea of model predictive control. These existing researches have explored various dispatching methods for EV charging. However, little literature has been able to consider the parking location of EV, such as buildings, in the scheduling. Additionally, only one trip instead of multiple trips for EVs is considered over the planning horizon in most of the literature. With the popularity of workplace and public charging to build range confidence, these will play an important role in improving the utilization of wind energy in supporting EV charging. Complementary to the existing researches, this paper focuses on the coordination of EV charging with locally generated wind power at multiple buildings, which incorporates the random driving requirements of EVs among the buildings. Meanwhile, multiple daily trips for each EV among the buildings are considered.

The challenges regarding this problem are as follows. *First*, there exist multiple uncertainties due to the volatile wind power supply of buildings and the random charging demand of EVs. *Second*, the charging decisions of EVs are not only constrained by their parking duration (to fulfill travel requirements) but also determined by their parking locations due to the time-dependent wind power supply at different buildings. *Third*, the problem is a multistage time-dependent decision process. Moreover, the charging decisions of EVs at the current stage affect the future cost. *Fourth*, the decision space of the problem increases exponentially with the number of EVs, thus the complexity to find an optimal charging policy is sensitive to the problem size.

In this paper, we first conduct a case study in Beijing to explore the potential of wind power of high-rise buildings in supporting EV charging. Considering that the charging demand of EVs usually doesn't align with the uncertain wind power generation of buildings, the coordination of EV charging with locally generated wind power in a microgrid of buildings is investigated and three main contributions are made. *First*, we investigate this problem and formulate it as a Markov decision process (MDP), which incorporates the uncertain wind power supply at different buildings and the random driving requirements of EVs among the buildings. *Second*, we develop a distributed simulation-based policy improvement (DSBPI) method as an extension of rollout [26], to improve from heuristic or experience-based policies while improving scalability. *Third*, the performance of the distributed policy improvement method is proved in this paper. Moreover, we compare DSBPI with a central version method on two case studies to evaluate the performance and scalability of DSBPI.

The remainder of this paper is outlined as follows. In Section II, a case study in Beijing on the wind power of high-rise buildings is conducted. In Section III, the problem is formulated as a MDP. In Section IV, a DSBPI method is proposed. In Section V, the DSBPI is evaluated on two case studies. In Section VI, the paper is briefly concluded.

II. The Potential of Building Mounted Wind Power in Beijing

In the literature, much effort has been made on the technical issues of building mounted wind turbines [7], [16]. However, little previous work has explored the potential of wind production of buildings, which directly determines its economic benefits of application. High-rise buildings are now pervasive in China, especially in big cities, such as Beijing. Considering that Beijing is more likely to firstly promote large adoption of EVs and widely develop wind power of buildings due to the rapidly developed technologies and the enormous high-rise buildings, we conduct a case study in the city to explore the potential of locally generated wind power of high-rise buildings in supporting EV charging in this section.

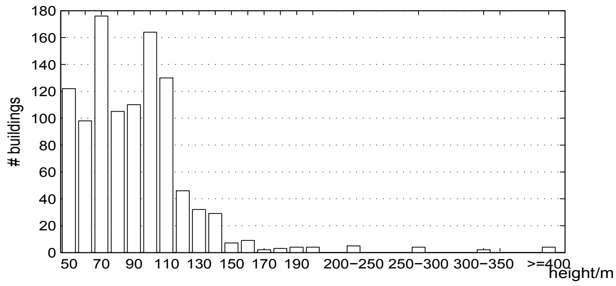


Fig. 1. The distribution of building height in Beijing

A. Statistics Data

Based on the incomplete report of EMPORIS [17], the height distribution for the recorded high-rise buildings ($\geq$ 50 m) in Beijing is plotted in histograms as Fig. 1. The statistics results reflect that the current population of buildings higher than 100 meters has exceeded 300 in the city. To investigate potential wind power production of buildings in Beijing, the monthly average wind speed of the city based on the observations at Beijing Capital Airport (heigh=55 m) from Jun. 2009 to Mar. 2016 [18] is listed in TABLE I.

TABLE I
MONTHLY AVERAGE WIND SPEED IN BEIJING

Month	Average Wind Speed (m/s)	Month	Average Wind Speed (m/s)
Jan.	4.1	Jul.	3.1
Feb.	3.6	Aug.	3.1
Mar.	4.1	Sep.	3.1
Apr.	4.6	Oct.	3.1
May	4.6	Nov.	3.6
Jun.	3.1	Dec.	4.1

B. Estimation Models

It has been acknowledged that it is desirable to explore wind power of high-rise buildings concerning the abundant wind resource. However, the wind speed of buildings at different heights is usually unavailable. In the literature, power law has been recognized as one common approach to estimate the wind speed at different heights [19], i.e.

$$u(z) = u(z_{\text{ref}}) \cdot (z/z_{\text{ref}})^{\alpha} \tag{1}$$

where $u(z)$ and $u(z_{\text{ref}})$ denote wind speed at height z and at reference height z_{ref}. The exponent α is an empirically derived coefficient dependent on terrain types. For city area with tall buildings, we have $\alpha = 0.4$ [19].

The instantaneous power output of wind turbines relates to a designated wind speed by a cubic law, i.e.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2707103, IEEE Transactions on Smart Grid

$$P_w(u) = 1/2 C_p \rho u^3 A \quad (2)$$

where C_p denotes the power coefficient of wind turbine, A is the rotor swept area, ρ is the air density.

Considering that the monthly average wind speed usually underestimates the annual wind production, commonly used Rayleigh models are derived in this paper to model the variation of wind speed in each month of a year [20], i.e.

$$f_{i,z}(u) = \frac{\pi u}{2\overline{u}_{i,z}^{\,2}} \exp(0.25\pi(\frac{u}{\overline{u}_{i,z}})^2), \forall i = 1, 2, ..., N_o \quad (3)$$

where i represents the month with $N_o = 12$. $f_{i,z}(\cdot)$ is probability density function for wind speed at height z in the i_{th} month. $\overline{u}_{i,z}$ denotes the average wind speed at height z of month i, which can be derived using (1).

According to [20], the annual average production of building mounted wind turbine at height z can be estimated as

$$\overline{P}_{w,z} = \frac{1}{N_o} \sum_{i=1}^{N_o} \int_{u_{\mathrm{ci}}}^{u_{\mathrm{co}}} P_w(u) \cdot f_{i,z}(u) du \quad (4)$$

where u_{ci} and u_{co} denotes the cutin and cutout wind speed of wind turbine.

C. Analysis Results

In the analysis, the power coefficient of wind turbines is set as $C_p = 0.45$, which refers to a newly designed wind turbine for buildings [8] with an estimated power coefficient of $C_p = 0.47$. The cutin and cutout wind speed is set as 3 m/s and 25 m/s based on the parameters of most commercial wind turbines.

- The reported daily average driving distance of private cars in Beijing is about 40.2 km [21], thus an annual average driving distance $14,700$ km. If a Kia Soul EV with a driving efficiency of 0.1617 kWh/km is considered [22], the annual average energy consumption of one private EV in Beijing is about $2,377$ kWh.
- We consider a special case, where one wind turbine of diameter 22 m is mounted on each high-rise building ($\geq$ 50 m) in Beijing. The estimated total annual wind production is about 1.69×10^8 kWh. It is enough to support about $71,000$ EVs, which is about three times of the current population of EVs in Beijing ($21,500$, at the beginning of 2016) [23]. We conclude that the estimated annual wind production of the recorded high-rise buildings can suffice more than three times of the EVs in the city.
- Besides, we consider a small-scale wind turbine, Aerostar 6 (diameter= 6.7 m, 10kW) in the case study. Suppose one wind turbine of that size is employed on each of those high-rise buildings. The estimated annual wind power production could reach 1.56×10^7 kWh, which is enough to suffice the travelling requirements of about $6,500$ EVs in the city. This implies that the annual production of four Aerostar 6 wind turbines on each of those buildings will be more than enough to support all the EVs in the city.

In the above case study, we consider two special cases and conclude that the potential wind production of high-rise buildings in Beijing is very promising. Specially, the annual production of a small number of wind turbines on those high-rise buildings is enough to potentially support the current population of EVs in the city. Besides, in practical deployment, if more wind power generators can be installed on the high-rise buildings, it is possible to suffice a larger population of EVs. However, the charging load of EVs parked in buildings is usually uncoordinated with the locally generated wind power of buildings due to their random driving requirements. Considering that EVs are usually parked in buildings for long time periods, the problem to schedule EV charging to align with locally generated wind power in a microgrid of buildings is investigated in this paper.

III. Problem Formulation

A. System Description

In this part, we consider a microgrid consisting of multiple high-rise buildings, on-site wind power of buildings and a number of EVs. The system architecture is depicted as Fig. 2. The buildings include residential, office and commercial buildings. The EVs are usually parked in the buildings or driven among the buildings. For example, an office worker may drive to work from a residential building in the morning and park the EV in an office building till after work. The driving behaviors of EVs are random, which depend on the travel requirements of EV owners. The wind power generated at different buildings is stochastic. In the system, the role of buildings with locally generated wind power is to offer charging service to the EV owners to make profits. And the EV owners are motivated to register as dispatching load by a lower charging cost offered by the buildings compared with getting charged in the state grid without participation. Meanwhile, the buildings are contractually obligated to suffice the EV owners' traveling requirements every day. To improve the utilization of wind power of buildings, we consider a more general case, where neighboring buildings (close to each other in distance) can share excess wind power with each other. The locally generated wind power of a building is first used to charge EVs parked there. When it's not enough to satisfy the EVs' charging demand, surplus wind power from its neighboring buildings or purchased electricity from the power grid can supplement.

In the system, there is a local coordinator corresponding to each building and one system operator. The EVs are required to report their charging demand (next trip time) and parking duration to the local coordinator upon arrival at a building. Meanwhile, the local coordinators are required to inform the system operator of local charging information as well as on-site wind power generation. Considering that the marginal cost of wind power is usually much lower than thermal generation, wind power is assumed free in the problem (the installation and maintenance fees are ignored). However, power supplement from the power grid is charged with time-of-use (TOU) tariffs.

In order to simplify the discussion, the charging power is set constant in this paper. The problem is formulated and discussed in a discrete time framework on a daily basis, where each day is equally discretized into $T = 48$ stages with a $\Delta t = 30$ minutes' interval.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2707103, IEEE Transactions on Smart Grid

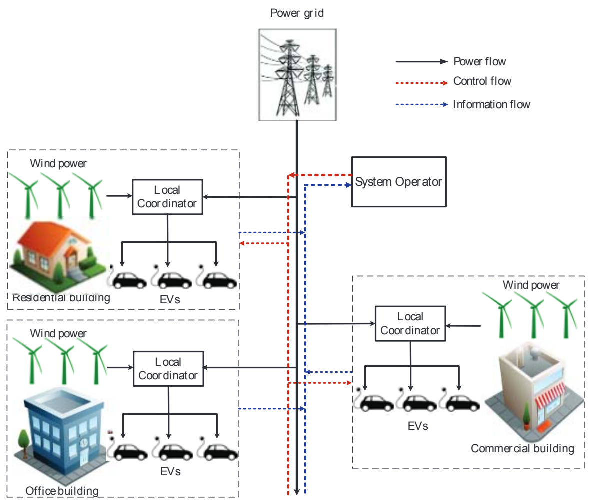


Fig. 2. System Architecture

B. Modeling

We consider a population of N EVs over the optimization horizons $\mathcal{T} \triangleq \{1,2,...,T\}$ in a microgrid of M buildings. The EVs and buildings are labelled as $\mathcal{N} \triangleq \{1,2,...,N\}$ and $\mathcal{M} \triangleq \{1,2,...,M\}$, respectively. The problem to coordinate EV charging with locally generated wind power of multiple buildings is modeled as a finite-stage MDP problem as follows.

*1) **System States**:* the system state at time t is defined as $S_t = [W_t^j, E_t^i, L_t^i, D_t^i]$ ($j \in \mathcal{M}$ and $i \in \mathcal{N}$), where W_t^j is the on-site wind power generation at building j. E_t^i represents the remaining required charging energy. L_t^i is the remaining parking time or remaining trip time. D_t^i denotes the location of EV i at time t. In the formulation, the location space is defined as $D_t^i \in \mathcal{D} \triangleq \{0,1,2,...,M\}$. When EV i is parked in building j ($j \in \mathcal{M}$), we define $D_t^i = j$, $E_t^i \geq 0$ and L_t^i as the remaining parking time. When EV i is on travel, we define $D_t^i = 0$, $E_t^i = 0$, and L_t^i as the remaining trip time.

*2) **Action Space**:* as mentioned, the charging power for EVs is constant, thus a binary vector $A_t = [z_t^1, z_t^2, ..., z_t^N]$ is used to denote the charging decisions at time t with $z_t^i \in \{0,1\}$. If $z_t^i = 1$, EV i is selected to be charged, otherwise $z_t^i = 0$.

*3) **System Dynamics**:* with the state S_t and charging decision A_t given, the dynamics of state are depicted as follows.

The dynamics of the remaining parking time ($D_t^i > 0$) or remaining trip time ($D_t^i = 0$) for EV i are described as

$$L_{t+1}^i = \begin{cases} L_t^i - \Delta t & \text{if } L_t^i > 0 \\ \tau^{t+1} & \text{if } L_t^i = 0, D_t^i = 0 \\ \eta^{t+1} & \text{if } L_t^i = 0, D_t^i > 0 \end{cases} \tag{5}$$

where τ^{t+1} and η^{t+1} are random variables, τ^{t+1} represents the parking time of EV i when it arrives at a building at time $t+1$, η^{t+1} denotes its trip time departure at time $t+1$.

The location transition of EV i is depicted as

$$D_{t+1}^i = \begin{cases} D_t^i & \text{if } L_t^i > 0 \\ R^{t+1} & \text{if } L_t^i = 0, D_t^i = 0 \\ M+1 & \text{if } L_t^i = 0, D_t^i > 0 \end{cases} \tag{6}$$

where $R^{t+1} \in \mathcal{M}$ is a random variable and denotes the building that EV i arrives at time $t+1$.

The transition of the remaining required charging energy of EV i is written as

$$E_{t+1}^i = \begin{cases} E_t^i - z_t^i \cdot P \cdot \Delta t & \text{if } L_t^i \geq 0, D_t^i > 0 \\ \max(B_{t+1}^i - f_i(\eta^{t+1}), 0) & \text{if } L_t^i = 0, D_t^i = 0 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where P represents the constant charging power of EVs, B_{t+1}^i denotes the energy state of EV i when it arrives at a building at time $t+1$. The function $f_i(\cdot)$ describes the relationship between energy consumption of EV i and its trip time. The energy consumption rate of an EV relates to velocity, acceleration and roadway grade. In our discussion, the energy consumption rate Q_i is assumed constant, thus $f_i(\eta^{t+1}) = Q_i \cdot \eta^{t+1}$ ($i \in \mathcal{N}$).

*4) **Constraints**:* in our formulation, the charging decision variable A_t is implicitly constrained as follows,

$$0 \leq E_t^i \leq E_{\text{cap}}^i, \ \forall i \in \mathcal{N}, t \in \mathcal{T} \tag{8}$$

$$0 \leq E_t^i \leq L_t^i \cdot P \cdot \Delta t, \ \forall i \in \mathcal{N}, t \in \mathcal{T} \tag{9}$$

$$z_t^i \leq D_t^i, \ \forall i \in \mathcal{N}, t \in \mathcal{T} \tag{10}$$

where E_{cap}^i denotes the battery capacity of EV i. The constraint (8) imposes that the remaining required charging energy can't exceed the EV battery capacity. The constraint (9) implies that the remaining required charging energy can't exceed the maximum possible charging energy during the remaining parking time. The constraint (10) guarantees that the EVs can only get charged during the parking durations.

*5) **Objective Function**:* in this part, we first classify the neighboring buildings (close in distance) in the microgrid into an aggregator. Among each building aggregator, excess wind power can be shared with each other for free. However, wind energy exchange among different building aggregators is not considered in this paper. We assume the buildings can be classified into K aggregators denoted by $\mathcal{K} \triangleq \{1,2,...,K\}$. And $\mathcal{M}^k$ is used to represent the collection of buildings in aggregator k, intuitively we have $\mathcal{M} = \cup_{k=1}^K \mathcal{M}^k$. Since wind energy generated at a building aggregator can be shared for free, the main cost for each building aggregator to offer charging service to EVs is to purchase electricity from the power grid to satisfy the surplus EV charging demands. In the scheduling, the buildings receive a fixed payment from each individual EV owner every day.Therefore the on-step profit for the buildings can be computed as follows.

$$\begin{aligned} \tilde{C}_t(S_t, A_t) = \sum_{i=1}^{N} m_i - \sum_{k=1}^{K} \Big(& \beta_t \max(P_t^k - W_t^k, 0)\Delta t \\ & - G(P_t^{T,k}) \Big) \end{aligned} \tag{11}$$

where m_i represents the fixed payment charged from EV owner i, which may be defined α ($0 \leq \alpha \leq 100$) percent lower than the minimum charging cost that the owner should pay to the state grid without participation as [24]. In this regard, the payment of each EV owner is dependent on the amount of energy charged to the vehicle. β_t represents the electricity price purchased from the state grid at time t. $P_t^k = \sum_{j \in \mathcal{M}^k} P_t^{\text{EV},j}$ with $P_t^{\text{EV},j} = \sum_{i \in \{i | D_t^i = j\}} z_t^i \cdot P$ denotes the total charging power of EVs parked in building aggregator

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2707103, IEEE Transactions on Smart Grid

k at time t. $W_t^k = \sum_{j \in \mathcal{M}^k} W_t^j$ denotes the total available wind power supply in building aggregator k. $P_t^{T,k}$ represents the energy transmission among building aggregator k, which can be computed as $P_t^{T,k} = \sum_{j \in \mathcal{M}^k} \max(P_t^{\mathrm{EV},j} - W_t^j, 0) - \max(P_t^k - W_t^k, 0)$. $G(\cdot)$ denotes the cost due to transmission loss among each building aggregatror. While the first term in (11) represents the revenue of the buildings charged from the EV owners, the second term is the cost for the buildings fulfill the EV owners' traveling requirements, and the third term represents the cost due to energy transmission loss.

Concerning the multiple randomness in the problem, the expected profit of the buildings can be calculated as

$$\tilde{J}(\pi, S_1) = \mathrm{E}^{\pi}_{S_1}\Big[\sum_{i=1}^{N} m_i - \sum_{t=1}^{T}\sum_{k=1}^{K}\{\beta_t \max(P_t^k - W_t^k, 0)\Delta t + G(P_t^{T,k})\}\Big] \tag{12}$$

where S_1 is the initial system state and $\pi = [A_1, A_2, ..., A_T]$ denotes the charging policy over the optimization horizons.

Therefore, the problem to maximize the profit of buildings can be described as problem **P1**, i.e.

$$\textbf{P1:} \quad \max \tilde{J}(\pi, S_1) \quad s.t \ (8)-(10) \tag{13}$$

From **P1**, we note that with a proper financial compensation (α) designed, the buildings are able to gain profits while reducing the cost for EV owners meanwhile. This has been demonstrated in [24]. Considering that the buildings are contractually obligated to charge all the EVs to fulfill their traveling requirements every day, the payment m_n for each individual EV owner is not presented in the objective function, therefore we define problem **P2** instead, i.e.

$$\textbf{P2:} \quad \min J(\pi, S_1) \quad s.t \ (8)-(10) \tag{14}$$

where $J(\pi, S_1) = \mathrm{E}^{\pi}_{S_1}[\sum_{t=1}^{T} \beta_t \sum_{k=1}^{K} \max(P_t^k - W_t^k, 0)\Delta t + G(P_t^{T,k})]$ denotes the total charging cost for the buildings to fulfill EV traveling requirements. Accordingly, the one-step cost is defined as $C_t(S_t, A_t) = \sum_{k=1}^{K}\{\beta_t \max(P_t^k - W_t^k, 0)\Delta t + G(P_t^{T,k})\}$. The intuitive interpretation of **P2** is that the profits of buildings can be maximized by minimizing the cost to supporting EV charging. To simplify the discussion, energy transmission loss is ignored in this paper.

As described above, the stochastic problem is formulated as a MDP. To solve the problem is to find a charging policy to minimize the total EV charging cost for the buildings. Though an optimal charging policy exists definitely, it is an intractable task to employ traditional dynamic programming method to find the solution due to the curses of dimensionality.

C. On-site Wind Power

The wind power generation is characterized by wind speed and the power curve of wind turbines. When the on-site wind speed is known, the wind power generated at building j can be calculated using the following equations [14].

$$W_t^j = \begin{cases} W_{\mathrm{cap}}^j, & v_{\mathrm{rated}} < v_t \le v_{\mathrm{cutout}} \\ W_{\mathrm{cap}}^j(\frac{v_t}{v_{\mathrm{rated}}})^3, & v_{\mathrm{cutin}} \le v_t \le v_{\mathrm{rated}} \\ 0, & \text{Otherwise} \end{cases} \tag{15}$$

where v_t is the on-site wind speed at time t. v_{rated}, v_{cutin}, v_{cutout} and W_{cap}^j represent the rated, cutin, cutout wind speed and wind power capacity for the wind turbines at building m.

IV. Distributed Method

Due to the difficulties to solve stochastic MDP problem, simulation-based policy evaluation is usually used to find an optimal charging policy. However, it is usually time-consuming due to the large policy space. In practice, there usually exist experience or heuristic charging policies, therefore a distributed simulation-based policy improvement (DSBPI) method is developed in this paper to improve from heuristic or experience-based policies as an extension of rollout [26]. The rollout algorithm is first designed to approximate the solution of sequential decision problems and later extended to MDP problems in [14] and [27]. In the distributed method, the original MDP problem is decomposed into subproblems corresponding to the building aggregators. Since the number of EVs parked in each building aggregator is often much less than the population of EVs in the microgrid, this method shows performance in improving scalability.

*1) **Sub-states**:* in the distributed method, the system state S_t is divided into sub-states with respect to the EVs parked in each building aggregator as $S_t = [S_t^1, S_t^2, ..., S_t^K, S_t^{K+1}]$. The first K sub-states correspond to K building aggregators and $K+1$ is related to the EVs on travel, i.e.

$$\forall k \in \mathcal{K},\ S_t^k = [W_t^j, E_t^i, L_t^i, D_t^i], j \in \mathcal{M}^k,$$
$$i \in \{i | \text{EV } i \text{ is parked in building aggregator } k \text{ at time } t\}$$
$$k = K+1,\ S_t^k = [E_t^i, L_t^i, D_t^i],$$
$$i \in \{i | \text{EV } i \text{ is on travel at time } t, \text{ i.e. } D_t^i = 0\}$$

*2) **DSBPI method**:* in the DSBPI method, we first assume there is a given base charging policy $\overline{\pi} = [\overline{\pi}^1, \overline{\pi}^2, ..., \overline{\pi}^{K+1}]$ for EVs parked in each building aggregator (e.g. a heuristic or experience-based policy). We define $\overline{\pi}^k = [\overline{d}_1^k, \overline{d}_2^k, ..., \overline{d}_T^k]$ as the base charging decisions for EVs parked in building aggregator k over the optimization horizons. It is straightforward that we have $\overline{d}_t^{K+1} = 0$ $(t = 1, 2, ..., T)$ (EVs on travel).

To develop a distributed policy improvement algorithm, the simulation-based policy improvement (SBPI) method in [27] can be presented in a distributed manner. We define a local Q-factors for building aggregator k ($\forall k \in \mathcal{K}$), which is used to evaluate the performance of sub-action d_t^k adopted by building aggregator k regarding the state $S_t = [S_t^k, S_t^{-k}]$ and the charging policy d_t^{-k} for the other building aggregators, i.e.

$$Q_t^k(S_t^k, d_t^k, S_t^{-k}, d_t^{-k}) = C_t(S_t^k, d_t^k, S_t^{-k}, d_t^{-k}) + \mathrm{E}[V(S_{t+1}^k, S_{t+1}^{-k})] \tag{16}$$

where S_t^{-k} , $\{S_t^n; n \in \mathcal{K} \cup \{K+1\}, n \neq k\}$ denotes the collection of sub-states except for building aggregator k. Accordingly, d_t^{-k} , $\{d_t^n; n \in \mathcal{K} \cup \{K+1\}, n \neq k\}$ is the collection of charging decisions for EVs related to the building aggregators except for building aggregator k. $V(S_{t+1}^k, S_{t+1}^{-k})$ is the optimal value function of building aggregator k, which depends on the optimal charging policy for EVs related to building aggregator k from time $t+1$ to T, i.e.

$$V(S_{t+1}^{k}, S_{t+1}^{-k}) = \mathrm{E}\left[\sum_{\tau=t+1}^{T} C_\tau(S_\tau^k, d_\tau^{*k}(S_\tau^k), S_\tau^{-k}, d_\tau^{-k})\right] \quad (17)$$

where d_τ^{*k} denotes the optimal charging decision for EVs parked in building aggregator k at time τ.

Afterwards, the optimal charging decision for EVs parked in building aggregator k at time t can be determined by comparing the local Q-factors, i.e.

$$d_t^{*k}(S_t^k) = \arg\min_{d_t^k \in \mathcal{A}_t^k} Q_t^k(S_t^k, d_t^k, S_t^{-k}, d_t^{-k}) \quad (18)$$

where $\mathcal{A}_t^k$ is decision space of EVs parked in building aggregator k at time t.

From above, we note that the local Q-factor for building aggregator k not only depends on the charging policy of EVs parked in the other building aggregators (d_t^{-k}) but also relies on the optimal charging policy for EVs parked in building aggregator k from time $t+1$. However, they are often unknown a priori before the problem is totally solved, thus the local Q-factors are intractable to accurately calculated.

To tackle this situation, the idea of policy improvement (also known as rollout) that using a base policy to estimate Q-factors is adopted in this paper. And a distributed policy improvement method is developed to improve from a given base policy while improving scalability. In the DSBPI method, the policy improvement algorithm is used by K building aggregators in succession. Consider a special case, in which the policy improvement is firstly used by building aggregator 1. We can use the base charging policy for the other building aggregators ($k \in \mathcal{K}\backslash\{1\}$) to approximate the local Q-factor for building aggregator 1. As a consequent, an improved charging policy $\hat{\pi}^1$ for building aggregator 1 can be obtained. After that the policy improvement algorithm is used by building aggregator 2, in which the improved policy $\hat{\pi}^1$ for building aggregator 1 and the base policy for the others ($k \in \mathcal{K}\backslash\{1,2\}$) are employed to estimate the local Q-factor. Likewise, an improved charging policy $\hat{\pi}^2$ for building aggregator 2 can be attained. The details of the method are described in **Algorithm** 1.

In the DSBPI method, when the policy improvement is used by building aggregator k, the local Q-factor is estimated as

$$\begin{aligned}\hat{Q}_t^k(S_t^k, d_t^k, S_t^{-k}, \bar{d}_t^{-k}) = C_t(S_t^k, d_t^k, S_t^{-k}, \bar{d}_t^{-k}) \\ + \mathrm{E}[\hat{V}(S_{t+1}^k, S_{t+1}^{-k})]\end{aligned} \quad (19)$$

with the approximate value function estimated as

$$\hat{V}(S_{t+1}^{k}, S_{t+1}^{-k}) = \mathrm{E}\left[\sum_{\tau=t+1}^{T} C_\tau(S_\tau^k, \bar{d}_\tau^{k}(S_\tau^k), S_\tau^{-k}, \bar{d}_\tau^{-k})\right] \quad (20)$$

with a little abuse of notation, we define $\bar{d}_t^{-k} = [\hat{d}_t^1, ..., \hat{d}_t^{k-1}, \bar{d}_t^{k+1}, ..., \bar{d}_t^{K+1}]$ in (19) and (20). It can be interpreted that the improved policy for building aggregators 1 to $k-1$ and the base policy for building aggregators $k+1$ to K are employed to estimate the local Q-factor for building aggregator k.

Besides, we note that there exist multiple randomness, making it computationally impossible to compute the expectations of the local optimal value function. Common random number technique is employed to approximate the expectations [14]. The main idea of common random number technique is a variance reduction method in which policy alternatives are evaluated against the common set of random sample paths. For instance, the local Q-factor at time t is estimated as

$$\begin{aligned}\hat{Q}_t^k(S_t^k, d_t^k, S_t^{-k}, \bar{d}_t^{-k}) = C_t(S_t^k, d_t^k, S_t^{-k}, \bar{d}_t^{-k}) \\ + \frac{1}{N_s}\sum_{\omega=1}^{N_s}\sum_{\tau=t+1}^{T}[C_\tau(S_\tau^k, \bar{d}_\tau^k(S_\tau^k), S_\tau^{-k}, \bar{d}_\tau^{-k})|\zeta_\omega]\end{aligned} \quad (21)$$

where N_s represents the number of sample paths and ζ_ω denotes the randomness of the sample path ω.

Based on the estimated local Q-factors, an improved charging decision for building aggregator k can be obtained, i.e.

$$\hat{d}_t^k(S_t^k) = \arg\min_{d_t^k \in \mathcal{A}_t^k} \hat{Q}_t^k(S_t^k, d_t^k, S_t^{-k}, \bar{d}_t^{-k}) \quad (22)$$

Considering that the charging decisions of EVs are binary variables (to charge or not to charge), the problem in (22) can be solved by enumeration for a small scale problem. However, for a large scale problem, the decision space is large, the technique of sampling or the reduction rules in [14] can be used to find a good enough solution.

In the literature [27], the performance of the SBPI method is proved that $J(\hat{\pi}, S_1) \leq J(\bar{\pi}, S_1), \forall S_1$ and $\bar{\pi}$, provided that the Q-factors are accurately estimated. In this paper, we show that this theorem still holds in the distributed manner. And the proof is attached in Appendix A.

Algorithm 1 The DSBPI method

Determine a base charging policy $\bar{\pi} = [\bar{d}_1, \bar{d}_2, ..., \bar{d}_T]$, where $\bar{d}_t = [\bar{d}_t^1, \bar{d}_t^2, ..., \bar{d}_t^{K+1}]$
Observe the initial state S_1 and divide it into sub-states as $S_1 = [S_1^1, S_1^2, ..., S_1^{K+1}]$
Generate N_s sample paths for on-site wind power at the buildings as well as parking events for the EVs.
for $k = 1, 2, ..., K$ **do**
    The SBPI method is used by building aggregator k
    for $t = 1, 2, ..., T$ **do**
        Divide the state S_t into sub-states as
$$S_t = [S_t^1, S_t^2, ..., S_t^{K+1}]$$
        Compute local Q-factor $\hat{Q}_t^k(S_t^k, d_t^k, S_t^{-k}, \bar{d}_t^{-k})$ as (21), where $\bar{d}_t^{-k} = [\hat{d}_t^1, ..., \hat{d}_t^{k-1}, \bar{d}_t^{k+1}, ..., \bar{d}_t^{K+1}]$
        Find $\hat{d}_t^k = \arg\min_{d_t^k \in \mathcal{A}_t^k} \hat{Q}_t^k(S_t^k, d_t^k, S_t^{-k}, \bar{d}_t^{-k})$
        Update the state $S_{t+1} = g(S_t, d_t)$, where $d_t = [\hat{d}_t^1, ..., \hat{d}_t^k, \bar{d}_t^{k+1}, ..., \bar{d}_t^{K+1}]$, $g(\cdot)$ denotes the system dynamics.
    end for
end for

V. Numeric Results

In this chapter, the performance and scalability of distributed simulation-based policy improvement(DSBPI) are evaluated

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2707103, IEEE Transactions on Smart Grid

through two case studies. Firstly, we conduct a comparative case study with $N = 12$ EVs, in which DSBPI and a central version method are compared. Afterwards, we give another case study with $N = 100$ EVs, in which the central version method is intractable, to show the scalability of DSBPI in coordinating the charging load of EVs with on-site wind power generated at buildings.

A. Parameters

We consider $M = 6$ high-rise buildings, including 2 residential buildings, 2 office buildings and 2 commercial building (shopping or entertainment center). In this case study, we assume the buildings are far way from each other except for the two commercial buildings, which are close in distance and can share excess wind power with each other. The major usage of EVs consists of commuting between home and workplace and travelling for shopping or entertainment. According to the research of Qi [29] and 2015 Beijing transportation report [31], most people in Beijing go to work between 7:00 am and 9:00 am and get off work between 5:00 pm and 7:00 pm. In [32], the stochastic characteristics of EV are analyzed based on real-world driving data and conclude that the departure and arrive time EVs show Gaussianlike distributions and correspond well to the general commuting times. Therefore the departure time of EVs from residential buildings and office buildings is described by a normal distribution $N(8{:}00\text{ am}, (1.0\text{h})^2)$ and $N(6{:}00\text{ pm}, (1.0\text{h})^2)$, respectively. Considering that the trip time of EVs depends on the distance, vehicle velocity and road congestion, normal distributions of different parameters are used to describe the variation of trip time between different buildings, which are listed in TABLE II. For instance, the trip time of EVs between residential building I and office building I obeys a normal distribution $N(1.0\text{h}, (0.5\text{h})^2)$. While the parking time of EVs in the residential and office buildings is featured by the random arrival and departure time, a normal distribution $N(1.2\text{h}, (1.2\text{h})^2)$ for EV parking duration in the commercial building is derived from the statistics analysis of Beijing private cars parked in commercial centers or supermarkets [25]. Without loss of generality, we assume the EV users drive to the commercial centers with a probability $p = 0.4$ or go home with a probability $(1-p) = 0.6$ after work in the case studies. The EV owners go to commercial I with a probability of $p = \frac{2}{3}$, and go to commercial II with a probability of $p = \frac{1}{3}$.

Considering that the EV users may live or work in different places, we use TABLE III to denote the percentages that live or work in the two residential or office buildings. For instance, 40% of EV users live in Residential I, among which 70% work in Office I and the remaining 30% work in Office II.

TABLE II
DISTRIBUTION OF TRIP TIME BETWEEN M = 5 BUILDINGS
Units:h (hour) $N(\mu, \sigma^2)$: normal distribution

	Office I	Office II	Commercial I (II)
Residential I	$N(1.0, (0.5)^2)$	$N(1.5, (0.5)^2)$	$N(1.0, (0.5)^2)$
Residential II	$N(0.5, (0.5)^2)$	$N(1.5, (0.5)^2)$	$N(1.5, (0.5)^2)$
Commercial I (II)	$N(1.0, (0.5)^2)$	$N(1.0, (0.5)^2)$	-

We assume the specification of the EVs are the same. The EV battery capacity is set as $E^i_{\text{cap}} = 60$ kWh, which refers to the parameters of BYD e6. The EV energy consumption rate is set as $Q_i = 5$ kW derived from the average driving speed of passenger cars in Beijing ($20 - 40$km/h) [33] and the electric energy efficiency of BYD e6 (0.2kWh/km). Besides, the constant charging power is set as $P = 4kW$. The parameters of wind turbines are selected based on the Ampair 10kW turbine ($v_{\text{cutin}} = 2.95$ m/s, $v_{\text{cutout}} = 50$ m/s and $v_{\text{rated}} = 11$ m/s).

TABLE III
DISTRIBUTION OF EV USERS IN THE RESIDENTIAL AND OFFICE BUILDINGS

Residence	Percentage	Working Place	Percentage
Residential I	40%	Office I	70%
		Office II	30%
Residential II	60%	Office I	40%
		Office II	60%

TABLE IV
THE TOU ELECTRICITY TARIFFS IN BEIJING

Type	TOU Price \$/kWh	Time Periods
Peak	0.138	8:00-12:00 and 17:00-21:00
Shoulder	0.109	12:00-17:00 and 21:00-24:00
Off-peak	0.058	0:00-8:00

A Rayleigh model with the annual average wind speed $\overline{u} = 3.6$ m/s is used to describe the stochastic wind speed in the numeric experiments. The industrial TOU electricity tariffs of Beijing [34] in TABLE IV are used as the electricity price from the power grid.

B. A Comparative Case Study

In this part, we consider a comparative case study with $N = 12$ EVs. The on-site wind power capacity for the buildings are set as 20kW (Residential I), 23kW (Residential II), 25kW (Office I), 20kW (Office II), 15kW (Commercial I), and 15kW (Commercial II). The DSBPI and a central version method are separately used to tackle the problem. The main idea of the central version method is to improve from heuristic or experienced-based policies of the original problem using SBPI without decomposition. The details can be found in [14], [28].

The base policy may be chosen based on experience-based rules or selected from some heuristic policies. Besides, there have been some existed works concerning how to choose from good candidates using efficient computing budget allocation, such as [30]. In this paper, we choose two heuristic policies as base policies. One is the greedy policy π_1 described in (23), in which the EVs begin to be charged upon arrival at a building until the required charging energy is achieved.

$$\pi_1 : z^i_t = \begin{cases} 1, & \text{if } D^i_t \leq M, L^i_t > 0 \text{ and } E^i_t > 0 \\ 0, & \text{Otherwise} \end{cases} \tag{23}$$

The other one is the myopic policy π_2 in (24), which seems a good candidate to reduce the total charging cost of EVs.

$$\pi_2 : \mathbf{z}_t = \arg\min_{\mathbf{z}_t \in \mathcal{A}_t} \sum_{k=1}^{M} |W^k_t - \sum_{i \in \{i | D^i_t \in \mathcal{M}^k\}} z^i_t \cdot P| \tag{24}$$

where $\mathbf{z}_t = [z^1_t, z^2_t, ..., z^N_t]$, $\mathcal{A}_t$ is decision space at time t.

Before the results are displayed, we define the notations.

- $\mathcal{I}_C(\cdot)$: improved policy using the central version method.
- $\mathcal{I}_D(\cdot)$: improved policy using the DSBPI method.

In this part, the DSBPI method and the central version method are evaluated in terms of the total charging cost of

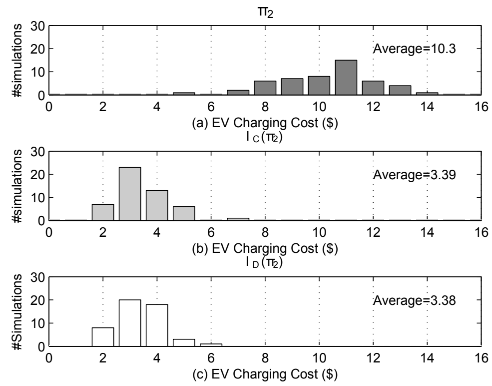


Fig. 3. The distribution of total charging cost of EVs of 50 sample paths with N = 12 EVs using (a) Base policy π_1. (b) Improved policy $\mathcal{I}_{\rm C}$ (π_1). (c) Improved policy $\mathcal{I}_{\rm D}$ (π_1).

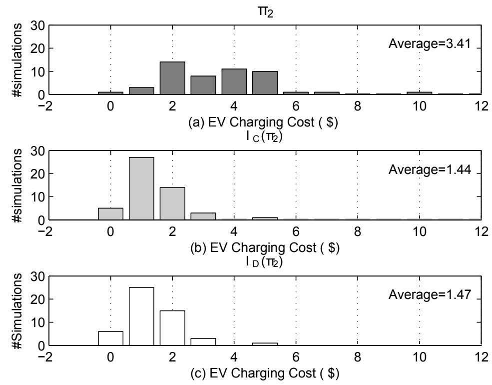


Fig. 4. The distribution of total charging cost of EVs of 50 sample paths with N = 12 EVs using (a) Basic policy π_2. (b) Improved policy $\mathcal{I}_{\rm C}$ (π_2). (c) Improved policy $\mathcal{I}_{\rm D}$ (π_2).

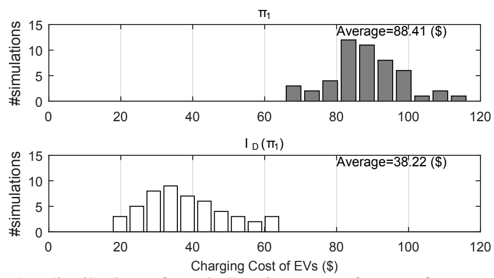


Fig. 5. The distribution of total charging cost of EVs of 50 sample paths with N = 100 EVs using (a) Base policy π_1. (b) Improved policy $\mathcal{I}_{\rm D}$ (π_1).

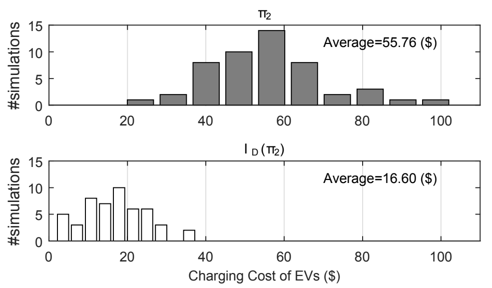


Fig. 6. The distribution of total charging cost of EVs of 50 sample paths with N = 100 EVs using (a) Base policy π_2. (b) Improved policy $\mathcal{I}_{\rm D}$ (π_2).

EVs defined in (12). A common set of $N_s = 50$ sample paths with different driving cycles for EVs and on-site wind power generation at the buildings is used to estimate Q-factors for the central version method and local Q-factors for DSBPI. The distributions of total charging cost of EVs resulting from different base policies and improved policies are plotted in histograms as Fig. 3 (base policy π_1) and Fig. 4 (base policy π_2). On average, the total charging costs of EVs using improved policies $\mathcal{I}_C(\pi_1)$ ($\mathcal{I}_C(\pi_2)$) and $\mathcal{I}_D(\pi_1)$ ($\mathcal{I}_D(\pi_2)$) are apparently reduced compared with base policies π_1 (π_2). Consequently, we conclude that the central version method and DSBPI indeed improve base policies. Besides, we observe that the distributions of total charging cost of EVs resulting from the improved polices $\mathcal{I}_C(\pi_1)$ ($\mathcal{I}_C(\pi_2)$) and $\mathcal{I}_D(\pi_1)$ ($\mathcal{I}_D(\pi_2)$) in Fig. 3 (Fig. 4) are similar. We can conclude that DSBPI can compete with the central version method in improving from base policies, however the former is advantageous in reducing curses of dimensionality. Moreover, the improved policies $\mathcal{I}_C(\pi_2)$ ($\mathcal{I}_D(\pi_2)$) outperform the improved policies $\mathcal{I}_C(\pi_1)$ ($\mathcal{I}_D(\pi_1)$). This difference is a direct result of the better performance of base policy π_2 compared with base policy π_1.

C. A 100-EV Case Study

In this part, we consider another case study with $N = 100$ EVs. The on-site wind power capacity for $M = 5$ buildings are set as 160kW (Residential I), 130kW (Residential II), 200kW (Office I), 160kW (Office II) and 100kW (Commercial). And the buildings are assumed far away form each other in this part. For the case, where some buildings are close to each other, the EVs parked in the same building aggregator can be aggregated in the algorithm to allow wind energy sharing among the neighboring buildings.

With the growth of EVs in the system, the decision space increases exponentially. It will be an intractable task to use the central version method to deal with the problem. Considering that the EVs parked in the same building may be aggregated, time aggregators are introduced in DSBPI. Specifically, the EVs parked in each building are aggregated by their remaining parking time at each stage. Thus, the subproblem related to a local coordinator is to decide the amount of power charged to the aggregators. Afterwards, one optimal decision to select a certain number of EVs to get charged in each time aggregator, can be directly decided by the remaining required charging energy in the descent order, which has been proved in [14]. Since the number of time aggregators is often much less than the number of EVs parked in each building, the decision space of subproblems in DSBPI will be greatly decreased.

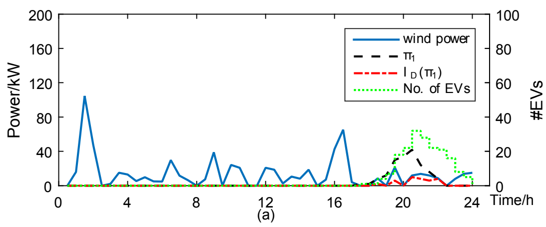


Fig. 9. On-site wind power, the number of EVs, and EV charging power using different policies in commercial building.

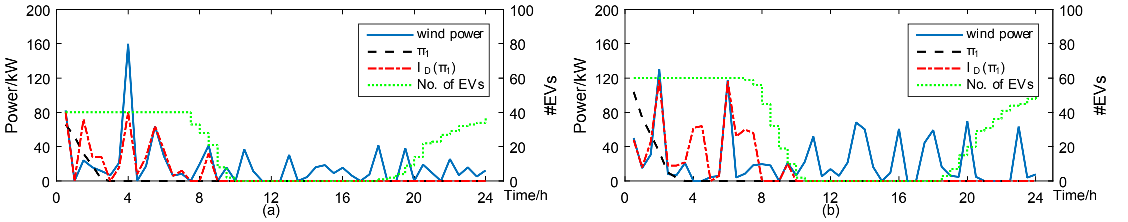


Fig. 7. On-site wind power, the number of EVs, and EV charging power using different policies in residential buildings: (a) Residential I. (b) Residential II.

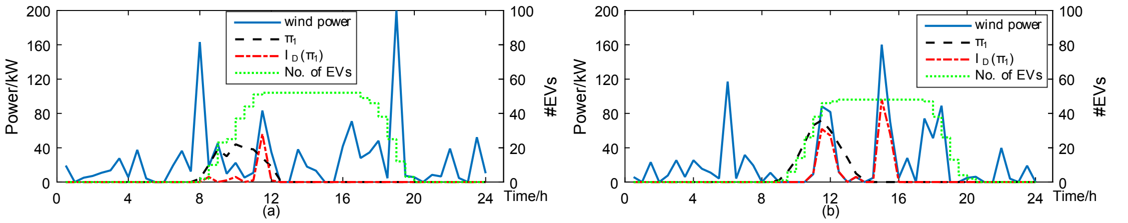


Fig. 8. On-site wind power, the number of EVs, and EV charging power using different policies in office buildings: (a) Office I. (b) Office II.

Similarly, $N_s = 50$ sample paths for the driving cycles for the EVs and wind power of the buildings are used to estimate the local Q-factors. The greedy policy π_1 and myopic policy π_2 are selected as two base policies. The distributions of total charging cost of EVs resulting from base policies and improved policies are plotted in histograms as Fig. 5 (base policy π_1) and Fig. 6 (base policy π_2). We can observe that the total charging costs of EVs are significantly reduced using improved policies $\mathcal{I}_D(\pi_1)$ ($\mathcal{I}_D(\pi_2)$) derived from DSBPI compared with base policies π_1 (π_2). While improved policy $\mathcal{I}_D(\pi_1)$ outperforms base policy π_1 with 56.8% less total EV charging cost, improved policy $\mathcal{I}_D(\pi_2)$ produces a reduction of 70.2% charging cost of EVs compared with base policy π_2.

Moreover, the amount of on-site wind power at different buildings utilized to charge the EVs resulting from different policies is investigated. Here base policy π_1 is illustrated, while a similar result can be obtained regarding base policy π_2. Fig. 7 to Fig. 9 show the on-site wind power, the number of EVs, and total EV charing power using different policies (π_1 and $\mathcal{I}_D(\pi_1)$) in different buildings of a sample path. We conclude that there are different types of charging load in different types of buildings (residential, office or commercial buildings). As Fig. 7 shows, a peak charging load appears during the night to the early morning (0:00 am-8:00 am). The result is caused by a massive percent of EVs parked in the residential buildings during that time period. In Fig. 8, we observe a massive charging load from 8:00 am to 6:00 pm in the office buildings. The phenomenon derives from a large proportion of EVs parked in the office buildings during the working hours. From 6:00 pm to 12:00 pm, a massive percentage of EVs are used for shopping or entertainment. As a result, a peak charging load occurs in the commercial building during that time period as Fig. 9. Besides, we conclude that the charging power of EVs better follows the stochastic wind power generation at each building with the improved policy $\mathcal{I}_D(\pi_1)$ from DSBPI employed compared with base policy π_1.

VI. Conclusion

In this paper, a prospective issue to utilize on-site wind power of multiple buildings to charge EVs is explored considering the following facts. The challenges associated with mounting and integrating wind turbines on buildings have been gradually overcame due to modern technology advances. The recent technical review study vividly captures the facts and shows that the wind resource has a great potential to be fully explored and developed in the urban environment [7]. In this paper, we first use a case study in Beijing to show that wind power generation of high-rise buildings can potentially support all the EVs in this city. Considering that the random EV charging demand is uncorrelated with renewable energy, we use a MDP to formulate the coordination of EV charging with on-site wind power of multiple buildings in a microgrid. Due to the difficulties to find an optimal charging policy, we develop a distributed simulation-based policy improvement (DSBPI) method to improve from heuristic or experience-based policies while improving scalability. Two case studies are conducted in which DSBPI is compared with a central version method. We conclude that DSBPI demonstrate good performance in reducing charging cost and scalability.

At the current stage, the base loads of buildings other than the EV charging demands are not considered. However, the model can be extended to use the on-site wind power of buildings to satisfy the local demand first and then use the surplus to charge the EVs. This will be our future work.

Note that exploring the structural property of the specific Markov decision processes may lead to more efficient solution methodologies, e.g., [35] and [36]. Event-based optimization may provide an alternative scalable solution methodologies for large-scale problems [37], [38], [39]. These are interesting directions to explore in the future.

Appendix A
Proof of DSBPI method

In DSPBI, we use $\bar{\pi} = [\bar{\pi}^1, \bar{\pi}^2, ..., \bar{\pi}^{K+1}]$ and $\hat{\pi} = [\hat{\pi}^1, \hat{\pi}^2, ..., \hat{\pi}^{K+1}]$ to denote the base and improved charging policy correspond to the building aggregators as defined previously. Accordingly, $\bar{d}_t^k$ and $\hat{d}_t^k$ denote the base and improved charging decisions for EVs in building aggregator k at time t.

When policy improvement is first used by building aggregator k, we have -2mm

$$\begin{aligned} &C_1(S_1^1,\hat{d}_1^1(S_1^1),S_1^{-1},\bar{d}_1^{-1}(S_1^{-1})) \\ &\quad + \mathrm{E}\Big[\sum_{t=2}^{T} C_t(S_t^1,\bar{d}_1^1(S_t^1),S_t^{-1},\bar{d}_1^{-1}(S_t^{-1}))\Big] \\ &\le C_1(S_1^1,\bar{d}_1^1(S_1^1),S_1^{-1},\bar{d}_1^{-1}(S_1^{-1})) \\ &\quad + \mathrm{E}\Big[\sum_{t=2}^{T} C_t(S_t^1,\bar{d}_t^1(S_t^1),S_t^{-1},\bar{d}_t^{-1}(S_t^{-1}))\Big] \end{aligned} \tag{25}$$

Similarly, we can imply that

$$\begin{aligned} &C_1(S_1^1,\hat{d}_1^1(S_1^1),S_1^{-1},\bar{d}_1(S_1^{-1})) \\ &\quad + \mathrm{E}\Big[\sum_{t=2}^{T} C_t(S_t^1,\hat{d}_t^1(S_t^1),S_t^{-1},\bar{d}_t(S_t^{-1})\Big] \\ &\le C_1(S_1^1,\hat{d}_1^1(S_1^1),S_1^{-1},\bar{d}_1(S_1^{-1})) \\ &\quad + \mathrm{E}\Big[\sum_{t=2}^{T'} C_t(S_t^1,\hat{d}_t^1(S_t^1),S_t^{-1},\bar{d}_t(S_t^{-1})) \\ &\quad + \sum_{t=T'}^{T} C_t(S_t^1,\bar{d}_t^1(S_t^1),S_t^{-1},\bar{d}_t^{-1}(S_t^{-1})\Big] \\ &\le C_1(S_1^1,\bar{d}_1^1(S_1^1),S_1^{-1},\bar{d}_1^{-1}(S_1^{-1})) \\ &\quad + \mathrm{E}\Big[\sum_{t=2}^{T} C_t(S_t^1,\bar{d}_t^1(S_t^1),S_t^{-1},\bar{d}_t^{-1}(S_t^{-1}))\Big] \end{aligned} \tag{26}$$

The left-hand side (LHS) of (26) is $J(\hat{\pi}^1,\bar{\pi}^{-1},S_1^1,S_1^{-1})$ and the right-hand side (RHS) equals $J(\bar{\pi}^1,\bar{\pi}^{-1},S_1^1,S_1^{-1})$, thus

$$J(\hat{\pi}^1,\bar{\pi}^{-1},S_1^1,S_1^{-1}) \le J(\bar{\pi}^1,\bar{\pi}^{-1},S_1^1,S_1^{-1}) \tag{27}$$

Similarly, we consider policy improvement is used by the k_{th} building aggregator. With a little buse of notation, we use $\bar{\pi}^{-k} = [\hat{\pi}^1,...,\hat{\pi}^{k-1},\bar{\pi}^k,...,\bar{\pi}^{K+1}]$ and $\bar{d}_t^{-k} = [\hat{d}_t^1,...,\hat{d}_t^{k-1},\bar{d}_t^k,...,\bar{d}_t^{K+1}]$ to represent "base" policy in the DSBPI to estimate the local Q-factors for the k_{th} building aggregator. Similarly to (25)-(26), we can conclude that

$$J(\hat{\pi}^k,\bar{\pi}^{-k},S_1^k,S_1^{-k}) \le J(\bar{\pi}^k,\bar{\pi}^{-k},S_1^k,S_1^{-k}) \tag{28}$$

Therefore, when the policy improvement method is used by the K_{th} building aggregators, we can derive that

$$J(\hat{\pi}^K,\bar{\pi}^{-K},S_1^k,S_1^{-k}) \le J(\bar{\pi}^K,\bar{\pi}^{-K},S_1^k,S_1^{-k}) \tag{29}$$

with the "base" policy $\bar{\pi}^{-K} = [\hat{\pi}^1,\hat{\pi}^2...,\hat{\pi}^{K-1}] = \hat{\pi}^{-K}$.

Thus we conclude that $J(\hat{\pi},S_1) \le J(\bar{\pi},S_1)$ in DSBPI.

References

[1] N. Leemput, F. Geth, J. V. Roy, A. Delnooz, J. Büscher, and J. Driesen, "Impact of Electric Vehicle On-Board Single-Phase Charging Strategies on a Flemish Residential Grid," *IEEE Trans. on Smart Grid*, vol. 5, no. 4, pp. 1815-1822, 2014.

[2] A. Ipakchi, and F. Albuyes, "Grid of the Future," *IEEE Power and Energy Magazine*, vol. 7, no. 2, pp. 52-62, 2009.

[3] J. Gabel, M. Carver, and M. Gerometta, "CTBUH Year in Review: Tall Trends of 2015, and Forecasts for 2016," *Council on Tall Buildings and Urban Habit*, 2015 Year in Review Report.

[4] N. Mithraratne, "Roof-top wind turbines for micro-generation in urban houses in New Zealand," *Energy and Buildings*, vol. 41, no. 10, pp. 1013-1018, 2009.

[5] Encraft, Warwick Wind Trials Project, 2009.

[6] L. Liu, and K. Y. Ip, "Investigation on the feasibility and enhanced methods of wind power utilization in high-rise buildings of Hong Kong," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 2, pp. 450-461, 2009.

[7] T. F. Ishugah, Y. Li, R. Z. Wang, and J. K. Kiplagat, "Advances in wind energy resource exploitation in urban environment: A review," *Renewable and Sustainable Energy Reviews*, vol. 37, pp. 613-626, 2014.

[8] Treehunger, Silent Rooftop Wind Turbines could Generte Half of a Household's Energy Needs, press, Jun. 2014 [Online]. Available: http://www.treehugger.com/wind-technology/silent-wind-turbines-could-generate-half-household-energy.html

[9] M. R. Islam, S. Mekhilef, and R. Saidur, "Progress and recent trends of wind energy technology," *Renewable and Sustainable Energy Reviews*, vol. 21, pp. 456-468, 2013.

[10] O. Sundström, and C. Binding,"Flexible Charging Optimization for Electric Vehicles Considering Distributed Grid Constraints," *IEEE Trans. on Smart Grid*, vol. 3, no. 1, pp. 26-37, 2012.

[11] S. Vandael, B. Claessens, M. Hommelberg, T. Holvoet, and G. Deconinck, "A Scalable Three-Step Aproach for Demand Side Management of Plug-in Hybrid Vehicles," *IEEE Trans. on Smart Grid*, vol. 4, no. 2, pp. 720-728, 2013.

[12] T. Wu, Q. Yang, Z. Bao, and W. Yan, "Coordinated Energy Dispatch in Microgrid with Wind Power Generation and Plug-in Electric Vehicles," *IEEE Trans. on Smart Grid*, vol. 4, no. 3, 2013.

[13] P. Kou, D. Liang, L. Gao, and F. Gao, "Stochastic Coordination of Plug-in Electric Vehicles and Wind Turbines in Microgrid: A Model Predictive Control Approach," *IEEE Trans. on Smart Grid*, vol. 7, no. 3, pp. 1537-1551, 2016.

[14] Q. Huang, Q. -S. Jia, Z. Qiu, X. Guan, and G. Deconinck, "Matching EV Charging Load With Uncertain Wind: A Simulation-Based Policy Improvement Approach," *IEEE Trans. on Smart Grid*, vol. 6, no. 3, pp. 1425-1433, 2015.

[15] Y. Guo, J. Xiong, S. Xu, and W. Su, "Two-Stage Economic Operation of Microgrid-Like Electric Vehicles Parking Deck," *IEEE Trans. on Smart Grid*, vol. 7, no. 3, pp. 1703-1712, 2016.

[16] I. Abohela, N. Hamza, and S. Dudek, "Effect of roof shape, wind direction, building height and urban configuration on the energy yield and positioning of roof-mounted wind turbines," *Renewable Energy*, vol. 50, pp. 1106-1118, 2013.

[17] EMPORIS report [Online]. Available: http://www.emporis.com/city/100214/beijing-china.

[18] Forecast and Reports of WindFinder [Online]. Available: https://www.windfinder.com/windstatistics/beijing_capital_airport.

[19] T. Yeh, and L. Wang, "A Study on Generators Capacity for Wind Turbines Under Various Tower Heights and Rated Wind Speeds Using Weibull Distribution," *IEEE Trans. on Energy Convers.*, vol. 23, no. 2, pp. 592-602, 2008.

[20] S. L. Walker, "Building mounted wind turbines and their suitability for the urban scale-A review of methods of estimating urban wind resource," *Energy and Buildings*, vol. 43, no. 8, pp. 1852-1862, 2011.

[21] H. Wang, L. Wu, C. Hou, and M. Ouyang, "A GPS-based research on driving range and patterns of private passenger vehicle in Beijing," *IEEE Electric Vehicle Symposium and Exhibition (EVS27)*, pp. 1-7, 2013.

[22] E. Loveday, "2015 Kia Soul EV Gets Official EPA Range Rating Of 93 Miles C City Range Will Blow Your Mind," InsideEVs.com, 2014.

[23] International Energy Agency, Global EV Outlook 2015; 2015:

[24] E. Walraven, and M. T. J. Spaan, "Planning Under Uncertainty for Aggregated Electric Vehicle Charging with Renewable Energy Supply," *Proceedings of the 22nd European Conference on Artificial Intelligence*, pp. 904-912, 2016.

[25] Z. Luo, Z. Hu, Y. Song, Z. Xu, and H. Lu, "Optimal Coordination of Plug-in Electric Vehicles in Power Grids With Cost-Benefit Analysis-Part II: A Case Study in China," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3556-3565, 2013.

[26] D. P. Bertsekas, J. N. Tsitsiklis, and C. Wu, "Rollout Algorithms for Combinatorial Optimization," *J. of Heuristics*, vol. 3, no. 3, pp. 245-262, 1997.

[27] Q. -S. Jia, J. Shen, Z. Xu, and X. Guan, "Simulation-Based Policy Improvement for Energy Management in Commercial Office Buildings," *IEEE Trans. on Smart Grid*, vol. 3, no. 4, pp. 2211-2223, 2012.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2017.2707103, IEEE Transactions on Smart Grid

[28] Y. Yang, Q. -S. Jia, X. Guan, "The Joint Scheduling of EV Charging Load with Building Mounted Wind Power Using Simulation-Based Policy Improvement," *2016 International Syposium on Flexble Automation*, pp. 165-170, 2016.

[29] W. Qi, Z. Xu, Z. M. Shen, Z. Hu, and Y. Song, "Hierarchical Coordinated Control of Plug-in Electric Vehicles Charging in Multifamily Dwellings," *IEEE Trans. on Smart Grid*, vol. 5, no. 3, pp. 1465-1475, 2014.

[30] Q. -S. Jia, " Efficient computing budget allocation for simulation-based policy improvement," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 2, pp. 342-352, 2012.

[31] Beijing Transportation Research Center, Beijing Annual Tranportation Report, 2015 [Online]. Available:http://www.bjtrc.org.cn/.

[32] T. K. Lee, Z. Bareket, T. Gordon, and Z. S. Filipi, "Stochastic modeling for studies of real-world PHEV usage: Driving schedule and daily temporal distributions," *IEEE Trans. Veh. Technol.,* vol. 61, no. 4, pp. 1493-1502, 2012.

[33] L. Huan, L. Chunyu, J. Lents, N. Davis, M. Osses, and N. Nikkila, "Beijing Vehilce Activity Study," *ISSRC*, La Habra, 2005.

[34] Z. Xu, Z. Hu, Y. Song, W. Zhao, and Y. Zhang, "Coordination of PEVs charging across multiple aggregators," *Applied Energy*, No. 136, pp. 582-589, 2014.

[35] Q.-S. Jia, L. Shi, Y. Mo, and B. Sinopoli, "On optimal partial broadcasting of wireless sensor networks for Kalman filtering," *IEEE Trans. Autom. Control,* vol. 57, no. 3, pp. 715-721, 2012.

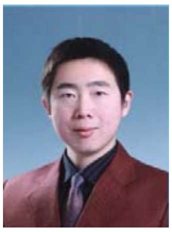
[36] Y. Zhao, X. Chen, Q.-S. Jia, X. Guan, S. Zhang, and Y. Jiang, "Long-term scheduling for cascaded hydro energy systems with annual water consumption and release constraints," *IEEE Trans. Autom. Sci. Eng.,* vol. 7, no. 4, pp. 969-976, 2010.

[37] Q.-S. Jia, "On solving optimal policies for finite-stage event-based optimization," *IEEE Trans. Autom. Control,* vol. 56, no. 9, pp. 2195-2200, 2011.

[38] Q.-S. Jia, "On solving event-based optimization with average reward over infinite stages," *IEEE Trans. Autom. Control,* vol. 56, no. 12, pp. 2912-2917, 2011.

[39] L. Xia, Q.-S. Jia, and X.-R. Cao, "A tutorial on event-based optimization - A new optimization framework," *Discrete Event Dynamic Systems - Theory and Applications,* vol. 24, no. 2, pp. 103-132, 2014.

Yu Yang received the B.E. degree in automation from Huazhong University of Science and Technology, Hubei, China, in 2013. She is currently pursuing the Ph.D. degree from the Center for Intelligent and Networked Systems (CFINS), Department of Automation, Tsinghua University, Beijing, China. Her current research interests include energy dispatching in smart grid, electric vehicle charging optimization, and decentralized optimization for large-scale dynamic systems.



Qing-Shan Jia received the B.E. and Ph.D. degrees in automation and control science and engineering from Tsinghua University, Beijing, China, in 2002 and 2006, respectively.

He was a Visiting Scholar at Harvard University, Cambridge, MA, USA, and the Hong Kong University of Science and Technology, Hong Kong, in 2006 and 2010, respectively. He is an Associate Professor with the Center for Intelligent and Networked Systems (CFINS), Department of Automation, TNList, Tsinghua University. His current research interests include theories and applications of discrete event dynamic systems, simulation-based performance evaluation, and optimization of complex systems.

Geert Deconinck received the M.Sc. degree in electrical engineering and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1991 and 1996, respectively.

He is full professor at the University of Leuven (KU Leuven, Belgium), where he heads the research group ELECTA on Electrical Energy in the Electrical Engineering Department. In the research centre EnergyVille on smart energy for sustainable cities, he is the scientific leader for the algorithms, modelling, optimisation, applied to smart electrical and thermal networks. His research focuses on robust distributed coordination and control, specifically in the context of smart grids. He is a fellow of the Institute of Engineering and Technology (IET) (M 2006, F 2013) and a senior member of IEEE (sM 1988, M 1996, SM 2000).

Xiaohong Guan received the B.S. and M.S. degrees in control engineering from Tsinghua University, Beijing, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical engineering from the University of Connecticut, Storrs, in 1993.

He was a Senior Consulting Engineer with PG&E from 1993 to 1995. He visited the Division of Engineering and Applied Science, Harvard University from January 1999 to February 2000. Since 1995, he has been with the Systems Engineering Institute, Xi'an Jiaotong University, Xi'an, China, and has been the appointed Cheung Kong Professor of Systems Engineering since 1999, and Dean of the School of Electronic and Information Engineering since 2008. Since 2001, he has also been the Director of the Center for Intelligent and Networked Systems, Tsinghua University, Beijing, China, and served as Head of the Department of Automation from 2003 to 2008. His research interests include complex networked systems including smart power grids, planning and scheduling of electrical power and manufacturing systems, and electric power markets.

Zhifeng Qiu received the M.S. degree in control engineering at Central South University in China and the Ph.D. degree in electrical engineering at KU Leuven in Belgium in 2004 and 2013 respectively.

Currently she is working in Central South University. Her research interests include techno-economic problem in power system, game-theoretic based analysis of power market and integration of renewable energy into smart building.

Zechun Hu received the B.S. degree and Ph.D. degree from Xian Jiao Tong University, Shaanxi, China, in 2000 and 2006, respectively.

He worked in Shanghai Jiao Tong University after graduation and also worked in University of Bath as a research officer from 2009 to 2010. He joined the Department of Electrical Engineering at Tsinghua University in 2010 where he is now an associate professor. He has published more than 120 papers and two book chapters. His major research interests include vehicle to grid techniques, applications of energy storage in power systems, optimal planning & operation of power systems, and power markets.